

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7258007号
(P7258007)

(45)発行日 令和5年4月14日(2023.4.14)

(24)登録日 令和5年4月6日(2023.4.6)

(51)国際特許分類	F I
G 1 0 L 15/28 (2013.01)	G 1 0 L 15/28 2 3 0 K
G 1 0 L 17/00 (2013.01)	G 1 0 L 17/00 2 0 0 C
G 0 6 F 3/16 (2006.01)	G 0 6 F 3/16 5 2 0

請求項の数 22 外国語出願 (全21頁)

(21)出願番号	特願2020-215271(P2020-215271)	(73)特許権者	521218881
(22)出願日	令和2年12月24日(2020.12.24)		オーナー デバイス カンパニー リミテッド
(62)分割の表示	特願2019-137200(P2019-137200)		中華人民共和国, 5 1 8 0 4 0, グアン
)の分割		ドン, シェンジェン, フーティアン デ
原出願日	平成27年4月10日(2015.4.10)		イストリクト, シェンミフ ストリート
(65)公開番号	特開2021-67951(P2021-67951A)		, ホンリ ウエスト ロード, ナンバー
(43)公開日	令和3年4月30日(2021.4.30)		8 0 8 9, シャム イップ スカイ パー
審査請求日	令和3年1月22日(2021.1.22)		ク, ビルディング 6, ユニット エー,
			スイート 3 4 0 1
		(74)代理人	100132481
			弁理士 赤澤 克豪
		(74)代理人	100115635
			弁理士 窪田 郁大
		(72)発明者	周 均 揚
			中華人民共和国 5 1 8 1 2 9 広東省深
			最終頁に続く

(54)【発明の名称】 音声認識方法、音声ウェイクアップ装置、音声認識装置、および端末

(57)【特許請求の範囲】

【請求項 1】

音声ウェイクアップ装置によって、周囲環境における音声情報を聴取するステップと、
聴取することによって取得された前記音声情報が音声ウェイクアップモデルと一致すると決定したとき、前記音声ウェイクアップ装置によって、聴取することによって取得される、第1のプリセットされた持続時間の音声情報をバッファし、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るステップであって、前記トリガ信号は、前記音声認識装置に対して、前記音声認識装置がイネーブルされた後、前記音声ウェイクアップ装置によってバッファされた前記音声情報を読み出し、認識するように指示するために用いられる、ステップと

を含む音声認識方法。

【請求項 2】

聴取することによって取得された前記音声情報が音声ウェイクアップモデルと一致すると決定する前記ステップは、

聴取することによって取得された前記音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された前記音声情報が前記音声ウェイクアップモデルと一致すると決定するステップを含む

請求項 1 に記載の方法。

【請求項 3】

聴取することによって取得された前記音声情報が音声ウェイクアップモデルと一致する

と決定する前記ステップは、

聴取することによって取得された前記音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴を抽出し、前記抽出された声紋特徴が所定の声紋特徴と一致すると決定し、聴取することによって取得された前記音声情報が前記音声ウェイクアップモデルと一致すると決定するステップを含む

請求項 1 に記載の方法。

【請求項 4】

前記声紋特徴は、ピッチ曲線、線形予測係数、スペクトル包絡線パラメータ、高調波エネルギー比率、共鳴ピーク周波数およびその帯域幅、ケプストラム、またはメル周波数ケプストラム係数などの前記声紋特徴を反映する音響パラメータを含む、

請求項 3 に記載の方法。

【請求項 5】

前記方法は、前のトリガ信号が受信された後の所定期間内に再び、さらなるトリガ信号が受信されなかったと決定するとき、前記音声認識装置を自動的にディセーブルすることをさらに含む、

請求項 1 ~ 4 のいずれか 1 項に記載の方法。

【請求項 6】

音声認識装置によって、音声ウェイクアップ装置によって送られたトリガ信号を受信するステップであって、前記トリガ信号は、前記音声認識装置に対して、それ自体をイネーブルし、前記音声ウェイクアップ装置によってバッファされた第 1 の音声情報を認識するように指示するために用いられる、ステップと、

前記トリガ信号を受信した後、前記音声認識装置によって、それ自体をイネーブル化し、第 2 のプリセットされた持続時間の第 2 の音声情報を聴取するステップと、

前記音声ウェイクアップ装置によってバッファされた前記第 1 の音声情報、および聴取することによって取得された前記第 2 の音声情報を認識して、認識結果を取得するステップと

を含む音声認識方法。

【請求項 7】

前記音声認識装置が前記認識結果を取得した後、前記方法は、

前記音声認識装置によって、前記取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うステップと、

前記音声認識装置によって、一致した音声命令情報に対応する動作を行うステップとをさらに含む

請求項 6 に記載の方法。

【請求項 8】

前記トリガ信号が受信された後、第 3 のプリセットされた持続時間内で前記トリガ信号が再び受信されなかったと決定したとき、前記音声認識装置によってそれ自体を自動的にディセーブルするステップをさらに含む

請求項 6 または 7 に記載の方法。

【請求項 9】

音声ウェイクアップ装置であって、

周囲環境における音声情報を聴取するように構成された聴取モジュールと、

前記聴取モジュールによって聴取することによって取得された前記音声情報が、音声ウェイクアップモデルと一致するかどうかを決定するように構成された決定モジュールと、

前記決定モジュールが、前記聴取モジュールによって聴取することによって取得された前記音声情報が前記音声ウェイクアップモデルと一致すると決定したとき、前記聴取モジュールによって聴取することによって取得される第 1 のプリセットされた持続時間の音声情報をバッファするように構成されたバッファモジュールと、

音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成された送出モジュールであって、前記トリガ信号は、前記音声認識装置に対して、前記音声認識装

10

20

30

40

50

置がイネーブルされた後、前記音声ウェイクアップ装置によってバッファされた前記音声情報を読み出し、認識するように指示するために用いられる、送出モジュールとを含む音声ウェイクアップ装置。

【請求項 10】

前記決定モジュールは、聴取することによって取得された前記音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された前記音声情報が前記音声ウェイクアップモデルと一致すると決定するように構成される

請求項 9 に記載の装置。

【請求項 11】

前記決定モジュールが、聴取することによって取得された前記音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声信号の声紋特徴を抽出するように構成された抽出モジュールをさらに備え、

前記決定モジュールは、前記抽出モジュールによって抽出された前記声紋特徴が所定の声紋特徴と一致すると決定したとき、聴取することによって取得された前記音声情報が前記音声ウェイクアップモデルと一致すると決定するようにさらに構成される

請求項 9 に記載の装置。

【請求項 12】

前記声紋特徴は、ピッチ曲線、線形予測係数、スペクトル包絡線パラメータ、高調波エネルギー比率、共鳴ピーク周波数およびその帯域幅、ケプストラム、またはメル周波数ケプストラム係数などの前記声紋特徴を反映する音響パラメータを含む、

請求項 11 に記載の装置。

【請求項 13】

前記音声認識装置は、前のトリガ信号が受信された後の所定期間内に再び、さらなるトリガ信号が受信されなかったと決定するとき、前記音声認識装置を自動的にディセーブルするように構成されたディセーブル化モジュールをさらに含む、

請求項 9 ~ 12 のいずれか 1 項に記載の装置。

【請求項 14】

音声認識装置であって、

音声ウェイクアップ装置によって送られたトリガ信号を受信するように構成された受信モジュールであって、前記トリガ信号は、前記音声認識装置に対して、それ自体をイネーブルし、前記音声ウェイクアップ装置によってバッファされた第 1 の音声情報を認識するように指示するために用いられる、受信モジュールと、

前記受信モジュールが前記トリガ信号を受信した後、それ自体をイネーブルし、第 2 のプリセットされた持続時間の第 2 の音声情報を聴取するように構成された聴取モジュールと、

前記音声ウェイクアップ装置によってバッファされた前記第 1 の音声情報、および前記聴取モジュールによって聴取することによって取得された前記第 2 の音声情報を認識して、認識結果を取得するように構成された認識モジュールと

を備える音声認識装置。

【請求項 15】

前記認識モジュールが認識を行った後に取得された前記認識結果と、予め記憶された音声命令情報との間の突き合わせを行うように構成された突き合わせモジュールと、

一致した音声命令情報に対応する動作を行うように構成された実行モジュールとをさらに備える

請求項 14 に記載の装置。

【請求項 16】

前記トリガ信号が受信された後、第 3 のプリセットされた持続時間内で前記トリガ信号が再び受信されなかったとき、前記認識モジュールをディセーブルするように構成されたディセーブル化モジュールをさらに備える

請求項 14 または 15 に記載の装置。

10

20

30

40

50

【請求項 17】

音声ウェイクアップ装置と、音声認識装置とを備え、

前記音声ウェイクアップ装置は、周囲環境における音声情報を聴取し、聴取することによって取得された前記音声情報が音声ウェイクアップモデルと一致すると決定したとき、第1のプリセットされた持続時間内で聴取することによって取得される第1の音声情報をバッファし、前記音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成され、

前記音声認識装置は、前記音声ウェイクアップ装置によって送られた前記トリガ信号を受信した後、それ自体をイネーブルし、第2のプリセットされた持続時間内で第2の音声情報を聴取し、前記音声ウェイクアップ装置によってバッファされた前記第1の音声情報、および聴取することによって取得された前記第2の音声情報を認識して認識結果を取得するように構成される端末。

10

【請求項 18】

前記音声ウェイクアップ装置は、デジタル信号プロセッサDSPである

請求項17に記載の端末。

【請求項 19】

前記音声認識装置は、アプリケーションプロセッサAPである

請求項17または18に記載の端末。

【請求項 20】

前記周囲環境における音声情報を聴取することは、前記端末が待機状態にあるかまたは画面ロック状態にあるときに周囲環境における音声情報を聴取することを含む、

20

請求項17～19のいずれか1項に記載の端末。

【請求項 21】

前記音声情報が音声ウェイクアップモデルに一致することは、

前記音声情報が、事前設定されたウェイクアップ音声情報に一致し、前記ウェイクアップ音声情報から抽出される声紋特徴が事前設定された声紋特徴に一致することを含み、前記声紋特徴は、ピッチ曲線、線形予測係数、スペクトル包絡線パラメータ、高調波エネルギー比率、共鳴ピーク周波数およびその帯域幅、ケプストラム、またはメル周波数ケプストラム係数などの前記声紋特徴を反映する音響パラメータを含む、

請求項17～20のいずれか1項に記載の端末。

30

【請求項 22】

前記端末は、前のトリガ信号が受信された後の所定期間内に再び、さらなるトリガ信号が受信されなかったと決定するとき、前記音声認識装置を自動的にディセーブルするようにさらに構成される、

請求項17～21のいずれか1項に記載の端末。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明はモバイル通信技術の分野に関し、詳細には、音声認識方法、音声ウェイクアップ装置、音声認識装置、および端末に関する。

40

【背景技術】**【0002】**

現在、モバイルハンドヘルド端末の人気の高まっており、特に携帯電話、タッチスクリーン技術もますます成熟してきている。タッチスクリーン技術は、ユーザの動作を容易にするが、呼び出し対話を完了するために複数のタッチステップが行われる必要があり、ユーザが運転している、またはタッチを行うのに都合が悪いとき、呼び出しをし損なうことがある。

【0003】

従って音声認識技術に基づいて呼び出しをすること、またはSMSメッセージを送ることなどの機能が出現している。さらに現在、新しく重要なユーザ対話技術として音声認識

50

は、モバイル端末にますます広く適用されている。

【0004】

しかし、音声認識技術に基づいて呼び出しをすること、またはSMSメッセージを送ることなどの現在のサービスは、タッチスクリーン技術が協調的に用いられるときのみ実施されることができる。

【発明の概要】

【0005】

本発明の実施形態は、音声認識方法、音声ウェイクアップ装置、音声認識装置、および端末を提供する。ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができる。ユーザはタッチスクリーンの助けを必要とせず、複数の命令を入力することも必要としない。

10

【0006】

第1の態様によれば本発明の実施形態は、音声認識方法を提供し、方法は、音声ウェイクアップ装置によって、周囲環境における音声情報を聴取するステップと、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声ウェイクアップ装置によって、聴取することによって取得される第1のプリセットされた持続時間の音声情報をバッファし、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るステップであって、トリガ信号は、音声認識装置に対して、音声認識装置がイネーブルされた後、音声ウェイクアップ装置によってバッファされた音声情報を読み出し、認識するように指示するために用いられる、ステップとを含む。

20

【0007】

第1の態様に関連して、第1の態様の第1の可能な実装形態において、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップは、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップを含む。

【0008】

第1の態様に関連して、第1の態様の第2の可能な実装形態において、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップは、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴を抽出し、抽出された声紋特徴が所定の声紋特徴と一致すると決定し、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップを含む。

30

【0009】

第2の態様に関連して、本発明の実施形態は、音声認識方法を提供し、方法は、音声認識装置によって、音声ウェイクアップ装置によって送られたトリガ信号を受信するステップであって、トリガ信号は、音声認識装置に対して、それ自体をイネーブルし、音声ウェイクアップ装置によってバッファされた第1の音声情報を認識するように指示するために用いられる、ステップと、

トリガ信号を受信した後、音声認識装置によって、それ自体をイネーブル化し、第2のプリセットされた持続時間の第2の音声情報を聴取するステップと、

40

音声ウェイクアップ装置によってバッファされた第1の音声情報、および聴取することによって取得された第2の音声情報を認識して、認識結果を取得するステップとを含む。

【0010】

第2の態様に関連して、第2の態様の第1の可能な実装形態において、音声認識装置が認識結果を取得した後、方法は、

音声認識装置によって、取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うステップと、

一致した音声命令情報に対応する動作を、音声認識装置によって行うステップとをさらに含む。

50

【 0 0 1 1 】

第2の態様、または第2の態様の第1の可能な実装形態に関連して、第2の態様の第2の可能な実装形態において方法は、

トリガ信号が受信された後、第3のプリセットされた持続時間内でトリガ信号が再び受信されなかったと決定したとき、音声認識装置によってそれ自体を自動的にディセーブルするステップをさらに含む。

【 0 0 1 2 】

第3の態様によれば、本発明の実施形態は、音声認識方法を提供し、方法は、

音声ウェイクアップ装置によって、周囲環境における音声情報を聴取するステップと、
聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声ウェイクアップ装置によって、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るステップとを含む。

10

【 0 0 1 3 】

第3の態様に関連して、第3の態様の第1の可能な実装形態において、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップは、
聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップを含む。

【 0 0 1 4 】

第3の態様に関連して、第3の態様の第2の可能な実装形態において、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップは、
聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴を抽出し、抽出された声紋特徴が所定の声紋特徴と一致すると決定し、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するステップを含む。

20

【 0 0 1 5 】

第4の態様によれば本発明の実施形態は、音声認識方法を提供し、方法は、

音声認識装置によって、音声ウェイクアップ装置によって送られたトリガ信号を受信するステップと、

音声認識装置によって、トリガ信号を受信した後、それ自体をイネーブル化し、ユーザに音声プロンプト命令を送るステップと、

音声認識装置によって、音声プロンプト命令に従ってユーザによって入力された音声信号を記録し、音声信号に対する認識を行って認識結果を取得するステップとを含む。

30

【 0 0 1 6 】

第5の態様によれば本発明の実施形態はさらに、音声ウェイクアップ装置を提供し、装置は、

周囲環境における音声情報を聴取するように構成された聴取モジュールと、

聴取モジュールによって聴取することによって取得された音声情報が、音声ウェイクアップモデルと一致するかどうかを決定するように構成された決定モジュールと、

決定モジュールが、聴取モジュールによって聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、聴取モジュールによって聴取することによって取得される第1のプリセットされた持続時間の音声情報をバッファするように構成されたバッファモジュールと、

40

音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成された送出モジュールであって、トリガ信号は、音声認識装置に対して、音声認識装置がイネーブルされた後、音声ウェイクアップ装置によってバッファされた音声情報を読み出し、認識するように指示するために用いられる、送出モジュールとを含む。

【 0 0 1 7 】

第5の態様に関連して、第5の態様の第1の可能な実装形態において、決定モジュールは、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致す

50

ると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

【0018】

第5の態様に関連して、第5の態様の第2の可能な実装形態において、装置は、決定モジュールが、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声信号の声紋特徴を抽出するように構成された抽出モジュールをさらに含み、

決定モジュールは、抽出モジュールによって抽出された声紋特徴が所定の声紋特徴と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するようにさらに構成される。

10

【0019】

第6の態様によれば本発明の実施形態は、音声認識装置を提供し、装置は、音声ウェイクアップ装置によって送られたトリガ信号を受信するように構成された受信モジュールであって、トリガ信号は、音声認識装置に対して、それ自体をイネーブルし、音声ウェイクアップ装置によってバッファされた第1の音声情報を認識するように指示するために用いられる、受信モジュールと、

受信モジュールがトリガ信号を受信した後、それ自体をイネーブルし、第2のプリセットされた持続時間の第2の音声情報を聴取するように構成された聴取モジュールと、

音声ウェイクアップ装置によってバッファされた第1の音声情報、および聴取モジュールによって聴取することによって取得された第2の音声情報を認識して、認識結果を取得するように構成された認識モジュールとを含む。

20

【0020】

第6の態様に関連して、第6の態様の第1の可能な実装形態において、この装置は、認識モジュールが認識を行った後に取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うように構成された突き合わせモジュールと、

一致した音声命令情報に対応する動作を行うように構成された実行モジュールとをさらに含む。

【0021】

第6の態様、または第6の態様の第1の可能な実装形態に関連して、第6の態様の第2の可能な実装形態において、装置は、

トリガ信号が受信された後、第3のプリセットされた持続時間内でトリガ信号が再び受信されなかったとき、音声認識モジュールをディセーブルするように構成されたディセーブル化モジュールをさらに含む。

30

【0022】

第7の態様によれば、本発明の実施形態は、音声ウェイクアップ装置を提供し、装置は、周囲環境における音声情報を聴取するように構成された聴取モジュールと、

聴取することによって取得された音声情報が音声ウェイクアップモデルと一致するかどうかを決定するように構成された決定モジュールと、

決定モジュールが、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成された送出モジュールとを含む。

40

【0023】

第7の態様に関連して、第7の態様の第1の可能な実装形態において、決定モジュールは、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

【0024】

第7の態様に関連して、第7の態様の第2の可能な実装形態において、装置は、決定モジュールが、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声信号の声紋特

50

徴を抽出するように構成された抽出モジュールをさらに含み、

決定モジュールは、抽出された声紋特徴が所定の声紋特徴と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

【0025】

第8の態様によれば本発明の実施形態は、音声認識装置を提供し、装置は、

音声ウェイクアップ装置によって送られたトリガ信号を受信するように構成された受信モジュールと、

受信モジュールがトリガ信号を受信した後、それ自体をイネーブルし、ユーザに音声プロンプト命令を送るように構成された送出モジュールと、

音声プロンプト命令に従ってユーザによって入力された音声信号を記録し、音声信号を認識して認識結果を取得するように構成された処理モジュールとを含む。

【0026】

第9の態様によれば、本発明の実施形態は、端末を提供し、端末は、

音声ウェイクアップ装置と、音声認識装置とを含み、

音声ウェイクアップ装置は、周囲環境における音声情報を聴取し、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、第1のプリセットされた持続時間内で聴取することによって取得される第1の音声情報をバッファし、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成され、

音声認識装置は、音声ウェイクアップ装置によって送られたトリガ信号を受信した後、それ自体をイネーブルし、第2のプリセットされた持続時間内で第2の音声情報を聴取し、音声ウェイクアップ装置によってバッファされた第1の音声情報、および聴取することによって取得された第2の音声情報を認識して認識結果を取得するように構成される。

【0027】

第9の態様に関連して、第9の態様の第1の可能な実装形態において、音声ウェイクアップ装置は、デジタル信号プロセッサDSPである。

【0028】

第9の態様、または第9の態様の第1の可能な実装形態に関連して、第9の態様の第2の可能な実装形態において、音声認識装置は、アプリケーションプロセッサAPである。

【0029】

本発明の実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ。さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに、音声ウェイクアップ装置は、聴取することによって取得された音声情報をバッファし、音声認識装置は、イネーブルされた後に第2の音声情報を聴取し、バッファされた第1の音声情報および第2の音声情報を認識し、それにより音声認識装置がウェイクアップされた後に音声情報の取得を開始したとき、音声認識装置がイネーブルされる前にユーザによって送られた部分的音声情報の喪失が回避されることができ。

【図面の簡単な説明】

【0030】

【図1】本発明の実施形態による端末の概略構成図である。

【図2】本発明の実施形態による音声認識方法のフローチャートである。

【図3】本発明の実施形態による音声認識方法の別のフローチャートである。

【図4】本発明の実施形態による音声認識方法のさらに別のフローチャートである。

【図5】本発明の実施形態による音声認識方法のまたさらに別のフローチャートである。

【図6】本発明の実施形態による音声ウェイクアップ装置の概略図である。

【図7】本発明の実施形態による音声認識装置の概略図である。

【図8】本発明の実施形態による音声ウェイクアップ装置の別の概略図である。

【図9】本発明の実施形態による音声認識装置の別の概略図である。

【図10】本発明の実施形態による音声認識方法の概略図である。

10

20

30

40

50

【発明を実施するための形態】**【0031】**

本発明の目的、技術的解決策、および利点をより明らかにするために、以下ではさらに本発明を添付の図面を参照して詳しくさらに述べる。明らかに、説明する実施形態は本発明の実施形態のすべてではなく単に一部である。本発明の実施形態に基づいて当業者によって創造的な努力なしに得られる他のすべての実施形態は、本発明の保護範囲に入るものである。

【0032】

本発明の実施形態は、音声認識方法、音声ウェイクアップ装置、音声認識装置、および端末を提供する。ユーザは1つの命令だけを送る必要があり、ユーザの要求は満たされることができ、ユーザはタッチスクリーンの助けを必要とせず、複数の命令を入力することも必要としない。方法および装置は、同じ発明の概念に基づく。方法および装置によって問題を解決するための原理は同様であるので、端末、装置、および方法の実装形態は互いに参照してよく、繰り返しの説明は提供されない。

10

【0033】

本発明の実施形態は、端末を提供する。図1に示されるように、端末は、音声ウェイクアップ装置101、および音声認識装置102を含む。

【0034】

音声ウェイクアップ装置101は、デジタル信号プロセッサ(Digital Signal Processor、略してDSP)を用いることによって実施され得る。音声認識装置102は、アプリケーションプロセッサ(Application Processor、略してAP)を用いることによって実施され得る。音声認識装置102はさらに、中央処理装置(Central Process Unit、略してCPU)を用いることによって実施され得る。

20

【0035】

音声ウェイクアップ装置101は、周囲環境における音声情報を聴取し、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、第1のプリセットされた持続時間内で聴取することによって取得される第1の音声情報をバッファし、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成される。

30

【0036】

音声認識装置102は、音声ウェイクアップ装置によって送られたトリガ信号を受信した後、それ自体をイネーブルし、第2のプリセットされた持続時間内で第2の音声情報を聴取し、音声ウェイクアップ装置によってバッファされた第1の音声情報、および聴取することによって取得された第2の音声情報を認識して、認識結果を取得するように構成される。

【0037】

任意選択で、トリガ信号が受信された後、第3のプリセットされた持続時間内でトリガ信号が再び受信されなかったと決定したとき、音声認識装置102はそれ自体を自動的にディセーブルする。

40

【0038】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ、さらに解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに音声ウェイクアップ装置は、聴取することによって取得された音声情報をバッファし、音声認識装置は、イネーブルされた後に第2の音声情報を聴取し、バッファされた第1の音声情報および第2の音声情報を認識し、それにより音声認識装置がウェイクアップされた後に音声情報の取得を開始したとき、音声認識装置がイネーブルされる前にユーザによって送られた部分的音声情報の喪失を回避することができる。

【0039】

50

具体的には、一般にDSPがトリガ信号をイニシエートした後、イネーブルされた後、APは記録を行うために記録チャンネルをイネーブルする。一般に、記録は、APがイネーブルされた後に開始される。しかし、この解決策において、APがイネーブルされる前に、DSPはウェイクアップ情報を受信したとき、記録およびバッファリングを行うことを開始する。イネーブルされた後に、APは、音声情報を取得するために記録を行い続け、次いでDSPバッファから読み出された音声情報、およびAPがイネーブルされた後に取得された音声情報を認識する。ワンショットシナリオでは、DSPをウェイクアップしてから、DSPによる命令を送出するまでには、時間差がある。APがイネーブルされた後に記録が行われる場合、APがイネーブルされた後の音声情報のみが記録されることができ、上記の時間差内の音声情報は失われる。しかし、DSPがウェイクアップされたときに記録が開始され、バッファされる場合、上記の時間差内の音声情報は取得されることが

10

【0040】

例えばウェイクアップワードを話し始める時点は t_0 、ウェイクアップワードを話し終わる時点は t_1 、コマンドワードを話し始める時点は t_2 、APをイネーブルする時点は t_3 である。バッファは、 t_0 から t_3 までの音声情報を含む。しかし、APがイネーブルされたときにのみ記録が行われる場合、 t_3 後の音声情報のみが記録されることができ、 t_0 から t_3 までの音声情報は記録されることができない。従って本発明のこの実施形態において提供される解決策では、ウェイクアップのために用いられる音声情報の後の音声情報は取得されることができ、音声情報の喪失は回避され、それによって音声認識を改善する。

20

【0041】

任意選択で、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するために、下記の方法が特に用いられることがある。

【0042】

第1の実装形態：

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声情報は、音声ウェイクアップモデルと一致する。

【0043】

ユーザは、前もってプロンプト、例えば「やあ、リトルE」に従って、音声ウェイクアップ装置にウェイクアップ音声情報を設定することができる。代替的に、ウェイクアップ音声情報は、工場渡しの時点で端末にプリセットされる。周囲環境における音声情報を検出するとき、音声ウェイクアップ装置は音声情報を、記憶されたウェイクアップ音声情報と比較する。2つが同じ場合、音声情報は音声ウェイクアップモデルと一致し、トリガ命令が音声認識装置102に送られ、または、2つが異なる場合、音声ウェイクアップ装置101は、現在の検出された音声情報を捨てて、検出および決定作業を行い続け得る。

30

【0044】

第2の実装形態：

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴が抽出され、抽出された声紋特徴は所定の声紋特徴と一致し、聴取することによって取得された音声情報は音声ウェイクアップモデルと一致する。

40

【0045】

声紋特徴は、下記の特徴の1つまたは複数を含む。

ピッチ曲線、線形予測係数、スペクトル包絡線パラメータ、高調波エネルギー比率、共鳴ピーク周波数およびその帯域幅、ケプストラム(パワーケプストラムとも呼ばれる)、またはメル周波数ケプストラム係数(Mel Frequency Cepstrum Coefficient、略してMFCC)などの声紋特徴を反映する音響パラメータ。この実施形態は、上記で述べられた声紋特徴パラメータに限定されない。

50

【 0 0 4 6 】

本発明のこの実施形態において、設定装置がさらに含まれ得る。ユーザの声紋特徴は予め抽出され、音声ウェイクアップ装置に記憶される。例えばユーザは、プロンプトに従って、音声情報を設定モジュールの内部に記録し、次いで声紋特徴を抽出し、抽出された声紋特徴を音声ウェイクアップ装置に記憶し得る。

【 0 0 4 7 】

本発明のこの実施形態において提供される解決策を用いることによって、声紋特徴がウェイクアップモデルに追加され、周囲環境におけるノイズ、および他のユーザの音声入力はフィルタ除去されることのできるようにし、音声ウェイクアップ装置は、ユーザに対して信頼性のあるセキュリティをもたらすことができるようにする。

10

【 0 0 4 8 】

任意選択で、音声認識装置が認識結果を取得した後、音声認識装置は、取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行い、音声認識装置は、一致した音声命令情報に対応する動作の実行を制御する。

【 0 0 4 9 】

音声命令情報は、音声認識装置に予め記憶される。音声認識装置は、複数の音声命令情報を含む。

【 0 0 5 0 】

本発明のこの実施形態において、音声命令情報に対応する動作を行う実行モジュールがさらに含まれ得る。音声認識装置は、音声命令情報に対応する動作を行う実行モジュールに、実行命令を送り得る。例えばスピーカ、発光装置などが含まれる。

20

【 0 0 5 1 】

例えば周囲環境における音声情報がウェイクアップモデルを満足することを検出したとき、音声ウェイクアップモジュールは、2秒などの第1のプリセットされた持続時間の第1の音声情報をバッファし、音声認識モジュールをトリガしてそれ自体をイネーブルして第2の音声情報を聴取し、次いでバッファされた第1の音声情報および第2の音声情報を認識し、認識結果を音声命令情報とをファジー的に比較して、音声情報が音声命令情報の1つと一致するかどうかを決定する。例えば音声命令情報は、「着信音を再生」または「MP3を再生」など、着信音またはMP3を再生するよう指示する音声命令情報を含み、または「あなたはどこ？」など、答えるように指示する音声命令情報、または「カメラフラッシュをオンする」など、カメラフラッシュを点灯させるように指示する音声命令情報を含む。

30

【 0 0 5 2 】

携帯電話などの端末デバイスは、本発明のこの実施形態において提供される解決策を用いることによってサーチされることがある。自宅では一般に、携帯電話はランダムに置かれ、使用に必要な場合に携帯電話を見つけるのに多少の時間がかかる。本発明のこの実施形態において提供される解決策を用いることによって、「こんにちは、リトルE、あなたはどこ？」が発声されることがある。従って、携帯電話の音声ウェイクアップモジュールは音声情報を検出し、音声情報と、音声ウェイクアップモデル（例えば、ウェイクアップ音声情報は「こんにちは、リトルE」である）との間の突き合わせを行う。音声情報が音声ウェイクアップモデルと一致したとき、音声情報はバッファ内にバッファされ、トリガ信号が音声認識モジュールに送られる。音声認識モジュールはそれ自体をイネーブルし、音声情報を聴取し始め、次いでバッファされた音声情報、および聴取することによって取得された音声情報を認識して、認識結果（テキスト結果は、「こんにちは、リトルE、あなたはどこ」である）を取得し、テキスト結果と音声命令情報との間の突き合わせを行う。例えば「こんにちは、リトルE、あなたはどこ」と一致する音声命令情報がMP3再生である場合、ユーザに促すようにMP3音楽が再生される。

40

【 0 0 5 3 】

本発明のこの実施形態において提供される解決策を用いることによって、呼び出しがなされることがある。携帯電話が待機状態にあるか画面ロック状態にあるかに関わらず、携

50

帯電話の音声ウェイクアップモジュールは常にイネーブルされた状態にあり、ユーザによって送られる音声情報は、例えば「こんにちは、リトルE、リトルAを呼び出す」を聴取することによって取得されることができるようになる。次いで呼び出しが直接なされることができ、いかなる他の動作も必要ない。

【0054】

本発明の実施形態はさらに、音声認識方法を提供する。図2に示されるように方法は下記を含む。

【0055】

ステップ201：音声ウェイクアップ装置は、周囲環境における音声情報を聴取する。

【0056】

ステップ202：聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声ウェイクアップ装置は、聴取することによって取得される第1のプリセットされた持続時間の音声情報をバッファし、音声認識装置のイネーブル化をトリガするためのトリガ信号を送り、トリガ信号は、音声認識装置に対して、音声認識装置がイネーブルされた後、音声ウェイクアップ装置によってバッファされた音声情報を読み出し、認識するように指示するために用いられる。

【0057】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ、さらに解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに、聴取することによって取得された音声情報はバッファされ、音声認識装置はイネーブルされ、音声認識装置は音声情報を聴取し、次いでバッファされた音声情報および聴取することによって取得された音声情報を認識し、音声認識装置がウェイクアップされた後に音声情報を取得し始めるときに部分的音声情報の喪失を回避されることができるようになり、音声認識が改善される。

【0058】

任意選択で、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するために、下記の方法が特に用いられることがある。

【0059】

第1の実装形態：

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声情報は、音声ウェイクアップモデルと一致する。

【0060】

第2の実装形態：

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴が抽出され、抽出された声紋特徴は所定の声紋特徴と一致し、聴取することによって取得された音声情報は音声ウェイクアップモデルと一致する。

【0061】

本発明の実施形態はさらに、音声認識方法を提供する。図3に示されるように、方法は下記を含む。

【0062】

ステップ301：音声認識装置は音声ウェイクアップ装置によって送られたトリガ信号を受信し、トリガ信号は、音声認識装置に対して、それ自体をイネーブルし、音声ウェイクアップ装置によってバッファされた第1の音声情報を認識するように指示するために用いられる。

【0063】

ステップ302：トリガ信号を受信した後、音声認識装置はイネーブルされ、音声認識装置は第2のプリセットされた持続時間の第2の音声情報を聴取する。

10

20

30

40

50

【 0 0 6 4 】

ステップ 3 0 3 : 音声ウェイクアップ装置によってバッファされた第 1 の音声情報、および聴取することによって取得された第 2 の音声情報を認識して、認識結果を取得する。

【 0 0 6 5 】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは 1 つの命令を送ることだけが必要であり、ユーザの要求は満たされることができる。さらに解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに聴取することによって取得された音声情報はバッファされ、音声認識装置はイネーブルされ、音声認識装置は音声情報を聴取し、次いでバッファされた音声情報および聴取することによって取得された音声情報を認識し、音声認識装置がウェイクアップされた後に音声情報を取得し始めるときに部分的音声情報の喪失を回避されることができるようにし、音声認識が改善される。

10

【 0 0 6 6 】

任意選択で、音声認識装置が認識結果を取得した後、方法は、音声認識装置によって、取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うステップと、

音声認識装置によって、一致した音声命令情報に対応する動作の実行を制御するステップとをさらに含む。

【 0 0 6 7 】

任意選択で、トリガ信号が受信された後、第 3 のプリセットされた持続時間内でトリガ信号が再び受信されなかったと決定したとき、音声認識装置はそれ自体を自動的にディセーブルする。

20

【 0 0 6 8 】

本発明の実施形態はさらに、音声認識方法を提供する。図 4 に示されるように方法は下記を含む。

【 0 0 6 9 】

ステップ 4 0 1 : 音声ウェイクアップ装置は、周囲環境における音声情報を聴取する。

【 0 0 7 0 】

ステップ 4 0 2 : 聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声ウェイクアップ装置は、音声認識装置のイネーブル化をトリガするためのトリガ信号を送る。

30

【 0 0 7 1 】

任意選択で、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するために、下記の方法が特に用いられることがある。

【 0 0 7 2 】

第 1 の実装形態 :

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声情報は、音声ウェイクアップモデルと一致する。

【 0 0 7 3 】

第 2 の実装形態 :

聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致するとき、聴取することによって取得された音声信号の声紋特徴が抽出され、抽出された声紋特徴は所定の声紋特徴と一致し、聴取することによって取得された音声情報は音声ウェイクアップモデルと一致する。

40

【 0 0 7 4 】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは 1 つの命令だけを送ることが必要であり、ユーザの要求は満たされることができる。さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに、聴取することによって取得された音声情報はバッファされ、音声認識装置はそれ自体

50

をイネーブルし、音声情報を聴取し、次いでバッファされた音声情報および聴取することによって取得された音声情報を認識し、音声認識装置がウェイクアップされた後に音声情報を取得し始めるときに部分的音声情報の喪失を回避されることができるようし、音声認識が改善される。

【0075】

本発明の実施形態はさらに、音声認識方法を提供する。図5に示されるように、方法は下記を含む。

【0076】

ステップ501：音声認識装置は、音声ウェイクアップ装置によって送られたトリガ信号を受信する。

【0077】

ステップ502：トリガ信号を受信した後、音声認識装置はそれ自体をイネーブルし、ユーザに音声プロンプト命令を送る。

【0078】

ステップ503：音声認識装置は、音声プロンプト命令に従ってユーザによって入力された音声信号を記録し、音声信号を認識して認識結果を取得する。

【0079】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ、さらに解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。

【0080】

任意選択で、音声認識装置が、音声信号を認識して認識結果を取得した後、方法は、音声認識装置によって、取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うステップと、

音声認識装置によって、一致した音声命令情報に対応する動作の実行を制御するステップとをさらに含む。

【0081】

本発明の実施形態はさらに、音声ウェイクアップ装置を提供する。図6に示されるように、装置は、

周囲環境における音声情報を聴取するように構成された聴取モジュール601と、
聴取モジュール601によって聴取することによって取得された音声情報が、音声ウェイクアップモデルと一致するかどうかを決定するように構成された決定モジュール602と、

決定モジュール602が、聴取モジュール601によって聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、聴取モジュール601によって聴取することによって取得される第1のプリセットされた持続時間の音声情報をバッファするように構成されたバッファモジュール603と、

音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成された送出モジュール604であって、トリガ信号は、音声認識装置に対して、音声認識装置がイネーブルされた後、音声ウェイクアップ装置によってバッファされた音声情報を読み出し、認識するように指示するために用いられる、送出モジュール604とを含む。

【0082】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ、さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに、聴取することによって取得された音声情報はバッファされ、音声認識装置はそれ自体をイネーブルし、音声情報を聴取し、次いでバッファされた音声情報および聴取することによって取得された音声情報を認識し、音声認識装置がウェイクアップされた後に音声情報を取得し始めるとき、部分的音声情報の喪失を回避されることができ、音声認識が改善される。

10

20

30

40

50

【0083】

任意選択で、決定モジュール602は、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

【0084】

任意選択で、装置は、

決定モジュール602が、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声信号の声紋特徴を抽出するように構成された抽出モジュールをさらに含み、

決定モジュール602は、抽出モジュールによって抽出された声紋特徴が所定の声紋特徴と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するようにさらに構成される。

10

【0085】

本発明の実施形態はさらに、音声認識装置を提供する。図7に示されるように、装置は、音声ウェイクアップ装置によって送られたトリガ信号を受信するように構成された受信モジュール701であって、トリガ信号は、音声認識装置に対して、それ自体をイネーブルし、音声ウェイクアップ装置によってバッファされた第1の音声情報を認識するように指示するために用いられる、受信モジュール701と、

受信モジュール701がトリガ信号を受信した後、それ自体をイネーブルし、第2のプリセットされた持続時間の第2の音声情報を聴取するように構成された聴取モジュール702と、

20

音声ウェイクアップ装置によってバッファされた第1の音声情報、および聴取モジュールによって聴取することによって取得された第2の音声情報を認識して、認識結果を取得するように構成された認識モジュール703とを含む。

【0086】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ。さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。

【0087】

任意選択で、装置は、

認識モジュール703が認識を行った後に取得された認識結果と、予め記憶された音声命令情報との間の突き合わせを行うように構成された突き合わせモジュールと、

一致した音声命令情報に対応する動作を行うように構成された実行モジュールとをさらに含む。

30

【0088】

任意選択で、装置は、

受信モジュールがトリガ信号を受信した後、第3のプリセットされた持続時間内でトリガ信号が再び受信されなかったとき、音声認識モジュールをディセーブルするように構成されたディセーブル化モジュールをさらに含む。

【0089】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができ。さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。さらに、聴取することによって取得された音声情報はバッファされ、音声認識装置はそれ自体をイネーブルし、音声情報を聴取し、次いでバッファされた音声情報および聴取することによって取得された音声情報を認識し、音声認識装置がウェイクアップされた後に音声情報を取得し始めるときに部分的音声情報の喪失を回避されることができ。さらに、音声認識が改善される。

40

【0090】

本発明の実施形態はさらに、音声ウェイクアップ装置を提供する。図8に示されるよう

50

に、装置は、

周囲環境における音声情報を聴取するように構成された聴取モジュール801と、
聴取することによって取得された音声情報が音声ウェイクアップモデルと一致するかどうかを決定するように構成された決定モジュール802と、

決定モジュール802が、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定したとき、音声認識装置のイネーブル化をトリガするためのトリガ信号を送るように構成された送出モジュール803とを含む。

【0091】

任意選択で決定モジュール802は、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

10

【0092】

任意選択で、装置は、

決定モジュール802が、聴取することによって取得された音声情報が所定のウェイクアップ音声情報と一致すると決定したとき、聴取することによって取得された音声信号の声紋特徴を抽出するように構成された抽出モジュールをさらに含み、

決定モジュール802は、抽出された声紋特徴が所定の声紋特徴と一致すると決定したとき、聴取することによって取得された音声情報が音声ウェイクアップモデルと一致すると決定するように特に構成される。

【0093】

本発明の実施形態は、音声認識装置を提供する。図9に示されるように、装置は、音声ウェイクアップ装置によって送られたトリガ信号を受信するように構成された受信モジュール901と、

20

受信モジュール901がトリガ信号を受信した後、それ自体をイネーブルし、ユーザに音声プロンプト命令を送るように構成された送出モジュール902と、

音声プロンプト命令に従ってユーザによって入力された音声信号を記録し、音声信号を認識して認識結果を取得するように構成された処理モジュール903とを含む。

【0094】

本発明のこの実施形態において提供される解決策を用いることによって、ユーザは1つの命令だけを送ることが必要であり、ユーザの要求は満たされることができる。さらに、解決策は、端末が待機状態にあるか非待機状態にあるかに関わらず適用可能である。

30

【0095】

以下では、図10に示されるように、ソフトウェア実装処理を参照して本発明の実施形態について具体的に説明する。

【0096】

ソフトウェアの観点から、音声認識モジュールは、ドライブレイヤ、オーディオハードウェア抽象化レイヤ(Audio HAL)、フレームワーク(Framework)レイヤ、音声認識エンジン(VA Service)、およびアプリケーション設定(Setting)に分割されることができる。

【0097】

P1. イベントを報告する。具体的には、ドライブレイヤは、DSPからトリガ信号を受信した後、トリガイベントをFrameworkに報告する。

40

【0098】

P2. イベントを報告する。具体的には、Audio HALは、上記のトリガイベントをVA Serviceに報告する。

【0099】

P3. パラメータを設定する。具体的には、それはバッファからデータを読み出すように設定される。

【0100】

P4. VA Serviceをイネーブルする。

50

【0101】

P5. VA Serviceは、Frameworkに記録イネーブル命令を送る。

【0102】

P6. Frameworkは、記録イネーブル命令を受信した後、オーディオデータ読み出し命令をAudio HALに送る。

【0103】

P7. Audio HALは、Frameworkによって送られたオーディオデータ読み出し命令を受信した後、Bufferデータに対する読み出しをイネーブルする。

【0104】

P8. Audio HALは、Bufferデータ取得命令をドライブに送り、それによりドライブはBufferデータ取得命令をDSPに送り、次いでDSPはBufferデータをドライブに送る。

10

【0105】

P9. ドライブは、受信されたBufferデータをVA Serviceに報告する。

【0106】

P10. VA Serviceは、Bufferデータおよび記録データに対して、認識処理を行う。

【0107】

P11. VA Serviceは、記録停止命令をFrameworkに送る。

【0108】

P12. Frameworkは、記録停止命令を受信した後、オーディオデータ読み出し停止命令をAudio HALに送る。

20

【0109】

P13. Audio HALは、Frameworkによって送られたオーディオデータ読み出し停止命令を受信した後、Bufferデータに対する読み出しをディセーブルする。

【0110】

P14. Audio HALは、Bufferデータ取得停止命令をドライブに送る。

【0111】

当業者は、本発明の実施形態が、方法、システム、またはコンピュータプログラム製品として提供されることがあることを理解するべきである。従って、本発明は、ハードウェアのみの実施形態、ソフトウェアのみの実施形態、またはソフトウェアおよびハードウェアの組み合わせによる実施形態の形式を用いることがある。さらに、本発明は、コンピュータ使用可能プログラムコードを含む1つまたは複数のコンピュータ使用可能記憶媒体(限定的ではないが、ディスクメモリ、CD-ROM、光メモリなどを含む)上において実施されるコンピュータプログラム製品の形式を用いることがある。

30

【0112】

本発明については、本発明の実施形態による方法、デバイス(システム)、およびコンピュータプログラム製品のフローチャートおよび/またはブロック図を参照して説明する。コンピュータプログラム命令は、フローチャートおよび/またはブロック図における各処理および/または各ブロック、ならびにフローチャートおよび/またはブロック図における処理および/またはブロックの組み合わせを実施するために用いられることがあることを理解されるべきである。これらのコンピュータプログラム命令は、汎用コンピュータ、専用コンピュータ、埋め込みプロセッサ、またはマシンを生成する他の任意のプログラマブルデータ処理デバイスに対して提供されてよく、それによりコンピュータまたは他の任意のプログラマブルデータ処理デバイスのコンピュータまたはプロセッサによって実行される命令は、フローチャートにおける1つまたは複数の処理、および/またはブロック図における1つまたは複数のブロックにおいて、特定の機能を実施するための装置を生成する。

40

【0113】

50

これらのコンピュータプログラム命令は、コンピュータまたは他の任意のプログラマブルデータ処理デバイスに特定の方法で動くように指示することができるコンピュータ可読メモリに記憶されてよく、したがってコンピュータ可読メモリに記憶された命令は、指示装置を含む人工物を生成する。指示装置は、フローチャートの1つまたは複数の処理、および/またはブロック図の1つまたは複数のブロックにおける特定の機能を実施する。

【0114】

これらのコンピュータプログラム命令は、コンピュータまたは他のプログラマブルデータ処理デバイス上にロードされてよく、したがって一連の動作およびステップがコンピュータまたは他のプログラマブルデバイス上で行われ、それによってコンピュータによって実施される処理を生成する。従って、コンピュータまたは他のプログラマブルデバイス上で実行される命令は、フローチャートの1つまたは複数の処理、および/またはブロック図の1つまたは複数のブロックにおける特定の機能を実施するためのステップを提供する。

10

【0115】

本発明のいくつかの実施形態が述べられたが、当業者は、基本的な発明概念を習得した後、これらの実施形態に変更および修正を行うことができる。従って、下記の特許請求の範囲は、実施形態、および本発明の範囲内に入るすべての変更および修正を包含すると解釈されるものである。

【0116】

明らかに、当業者は本発明の実施形態の趣旨および範囲から逸脱せずに、本発明の実施形態に様々な修正および変形を行うことができる。本発明はこれらの修正および変形を、それらが以下の特許請求の範囲およびそれらの等価な技術によって定義される保護の範囲内に入ることを条件として包含するものである。

20

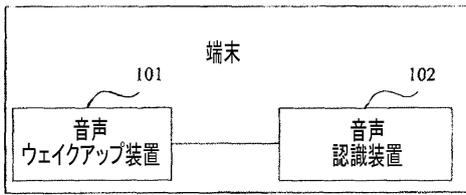
30

40

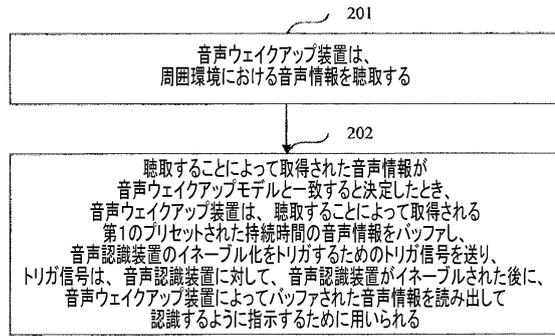
50

【 図 面 】

【 図 1 】

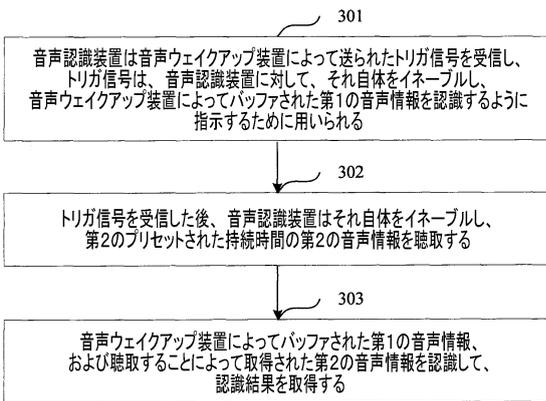


【 図 2 】

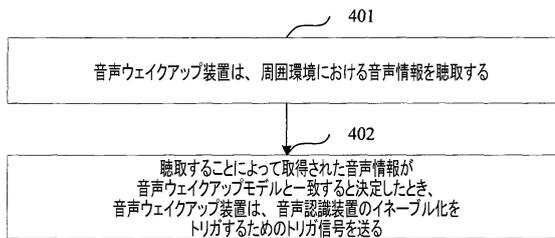


10

【 図 3 】

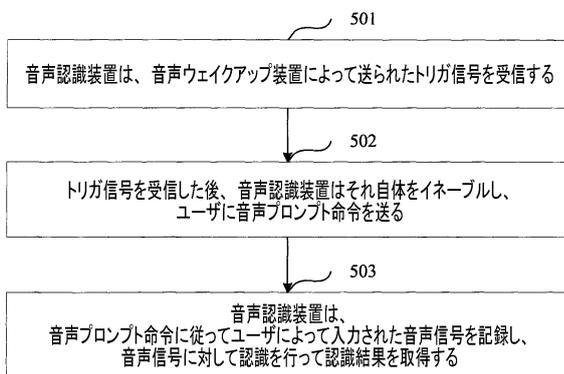


【 図 4 】

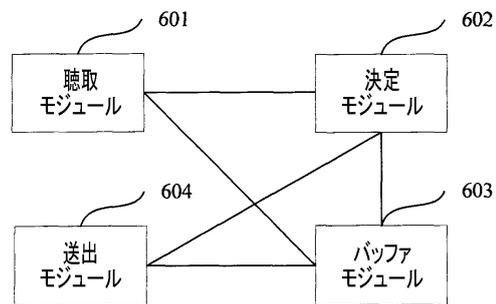


20

【 図 5 】



【 図 6 】

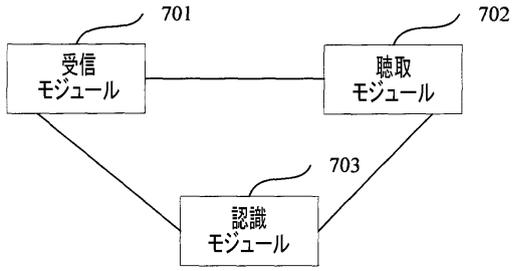


30

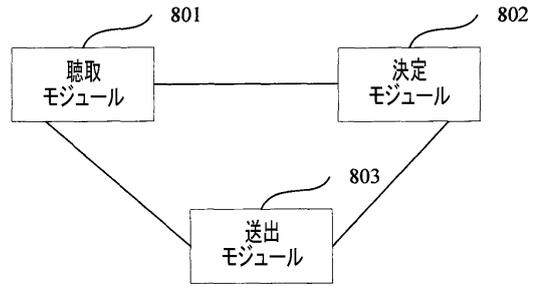
40

50

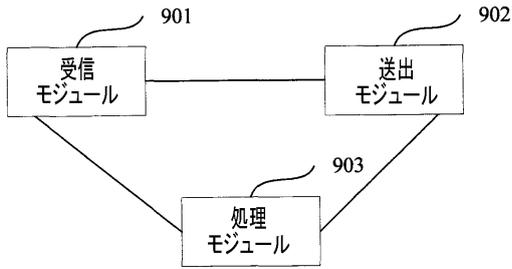
【図 7】



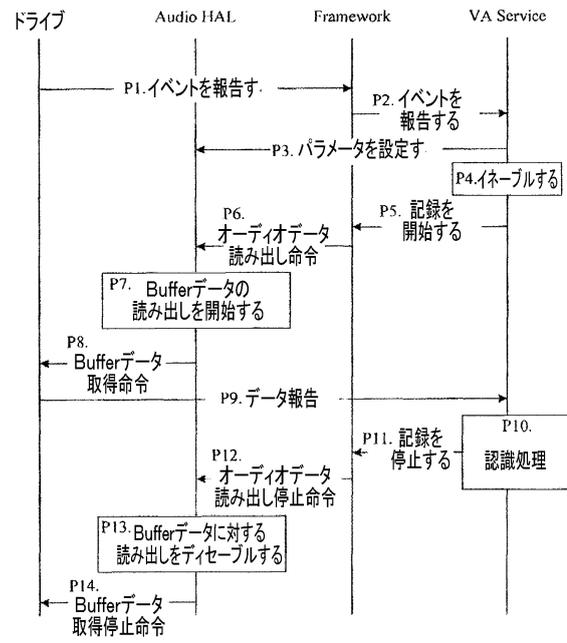
【図 8】



【図 9】



【図 10】



10

20

30

40

50

フロントページの続き

チェン 市龍崗区坂田 華為總部 ベン 公樓

審査官 中村 天真

- (56)参考文献 特表2017-520012(JP,A)
特開2015-052945(JP,A)
特開2015-011170(JP,A)
国際公開第2015/047932(WO,A1)
米国特許出願公開第2014/0222436(US,A1)
- (58)調査した分野 (Int.Cl., DB名)
G10L 15/00 - 17/26
G06F 3/01
G06F 3/048 - 3/0489
G06F 3/16
IEEE Xplore